AI and Blockchain for SOC Management: A Systematic Review and the CDSOCMP Framework

1st Ali Hamza
Department of Information and Communication Engineering
The Islamia University of Bahawalpur
Bahawalpur, Pakistan
ranaalihamza120@gmail.com

2nd Aziz ur Rehman
Department of Information and Communication Engineering
The Islamia University of Bahawalpur
Bahawalpur, Pakistan
azizyash793@gmail.com

3rd Muhammad Amjad
Department of Information and Communication Engineering
The Islamia University of Bahawalpur
Bahawalpur, Pakistan
amjadghafoor477@iub.edu.pk

4th Asjad Amin
Department of Information and Communication Engineering
The Islamia University of Bahawalpur
Bahawalpur, Pakistan
asjad.amin@iub.edu.pk

***Abstract*—Security Operations Centres (SOCs) face four persistent challenges: fragmented visibility across tools, high false-positive alert volumes, slow incident response, and forensic logs that remain vulnerable to tampering. This paper presents a systematic literature review of 97 studies published between 2013 and 2026, examining how artificial intelligence (AI) and blockchain technologies have been applied to these problems individually and in combination. The review follows a PRISMA-style selection process and includes a study-quality assessment of the included works. Based on the gaps identified in the literature, we propose the Cyber Defence Security Operations Centre Management Platform (CDSOCMP), a conceptual four-phase architecture intended to integrate AI-based detection, real-time blockchain-anchored forensics, automated response orchestration, and unified multi-cloud monitoring. We emphasise that CDSOCMP is presented as a conceptual architecture derived from the synthesis of existing literature; it has not yet been implemented or empirically evaluated, and no performance figures for CDSOCMP itself are reported. Reported accuracy and improvement figures discussed in this paper are drawn from the reviewed primary studies and are summarised here for comparative context only, not as results of CDSOCMP. The paper concludes with a discussion of how CDSOCMP relates to commercial SOC platforms, the threat of adversarial attacks on AI-driven defences, and a roadmap for prototype development and evaluation.**

***Index Terms*—Security Operations Centre (SOC), Artificial Intelligence, Blockchain, Automated Security Orchestration, Digital Forensics, Threat Detection, Multi-Cloud Security, Zero Trust Architecture**

I. Introduction

A. Background and Motivation

The cybersecurity landscape has changed substantially over the past decade [21], [22]. Organisations now face a steadily growing volume of attacks, and Security Operations Centres (SOCs)—the teams responsible for detecting, monitoring, and responding to security incidents—are increasingly struggling to keep pace [23]–[25], [27]. The literature repeatedly identifies four interrelated problems that continue to limit the effectiveness of modern SOCs:

- **Fragmented visibility:** Security tools frequently operate in isolation, leaving gaps across the attack surface that attackers can exploit [28], [29].
- **Alert fatigue:** A large proportion of alerts are false positives, forcing analysts to sift through thousands of alerts daily, which increases the risk of missed threats and analyst burnout [13], [31].
- **Slow incident response:** Manual remediation can take hours, while a modern automated attack can compromise an entire network within seconds [32], [33].
- **Forensic vulnerability:** Logging infrastructure in many legacy deployments can be altered after the fact, a weakness that skilled attackers and malicious insiders can exploit to erase evidence of their activity [34], [35].

To contextualise these issues, this section traces how SOC technology has evolved, broadly, across three overlapping generations (illustrated later in Fig. 5):

- *First generation (approx. 2000–2015):* Rule-based Security Information and Event Management (SIEM) platforms, with the bulk of analysis and response performed manually.
- *Second generation (approx. 2015–2022):* AI-assisted detection began to reduce manual triage effort, though analysts remained central to investigation and response.
- *Third generation (approx. 2022–present):* Research increasingly explores highly automated SOC concepts that integrate AI, orchestration, automation, and blockchain into a more unified operation. We note that fully autonomous SOCs of this kind remain largely at the research and pilot stage rather than being widely deployed in production.

DOI: https://doi.org/10.67535/tsp.000003.033
Published under the ICACNC 2026 Proceedings

B. The Research Gap

Considerable progress has been made on each of these technologies individually. However, the reviewed literature suggests that no single, empirically validated, integrated solution addresses all four SOC problems simultaneously. Specifically:

- AI-powered detection tools are often not tightly integrated with automated response systems.
- Where blockchain is used for security logging, it is typically applied *after* an event has been recorded, rather than as part of a real-time correlation pipeline.
- Few studies provide a genuinely unified view of security posture across multiple cloud providers.
- Forensic systems generally focus on collecting evidence *after* a breach, rather than continuously building a tamper-evident evidentiary record.

This paper aims to characterise these gaps through a systematic review, and to propose—at a conceptual level—an architecture, which we term CDSOCMP, intended to connect these four elements within a single design. We stress that the contribution of this paper is twofold and should not be conflated: (1) a systematic synthesis of existing peer-reviewed evidence (Sections III and V), and (2) a conceptual architectural proposal (Section IV) that is informed by, but distinct from, that evidence base. No claims are made that CDSOCMP has been built or tested.

II. Survey Methodology

A. Literature Search Strategy

We followed a process aligned with PRISMA (Preferred Reporting Items for Systematic Reviews and Meta-Analyses) guidelines. The search was conducted across five major databases: IEEE Xplore, ACM Digital Library, SpringerLink, ScienceDirect, and arXiv. Table I summarises the search parameters.

TABLE I
Literature Search Parameters

Parameter	Details
Databases	IEEE Xplore, ACM Digital Library, SpringerLink, ScienceDirect, arXiv
Timeframe	January 2013 – March 2026
Primary keywords	"Security Operations Centre", "SOC automation", "AI threat detection", "blockchain forensics"
Secondary keywords	"SIEM", "SOAR", "multi-cloud security", "federated learning security", "autonomous cyber defence"
Language	English only
Initial results	341 candidate papers
Final included	97 papers

B. Inclusion and Exclusion Criteria

A paper was **included** if it satisfied all of the following: (a) it addressed at least one of SOC operations, AI/ML-based threat detection, blockchain security, automated incident response, or multi-cloud monitoring; (b) it was published in a peer-reviewed venue or a recognised preprint repository; and (c) it reported empirical results, a described architecture, or a systematic analysis relevant to the review's scope.

A paper was **excluded** if it was: a duplicate or superseded version of an already-included study; outside the technical scope of this review (e.g., purely physical security or non-digital forensics); or judged, during full-text review, to have insufficient methodological rigour or validation to support its claims.

C. Screening, Selection, and Quality Assessment

Selection proceeded in three phases, illustrated in Fig. 1. Of the 341 papers initially identified, title and abstract screening removed 163 papers that were out of scope, leaving 178 papers for full-text review. Full-text review against the inclusion criteria above removed a further 66 papers for insufficient methodological rigour, leaving 112 papers. A subsequent quality check removed 15 additional papers due to weak experimental validation (e.g., absence of a described evaluation dataset, no comparison baseline, or internally inconsistent results), yielding a final corpus of 97 papers.

For the quality check, each candidate paper was assessed against four criteria, scored on a 0–2 scale (0 = absent/unclear, 1 = partially addressed, 2 = clearly addressed): (i) clarity of the stated research question and methodology; (ii) presence of an evaluation dataset or testbed; (iii) presence of a comparison baseline or ablation; and (iv) reproducibility of reported results (e.g., availability of parameters, dataset descriptions, or code). Papers scoring 4 or below out of 8 were excluded at this stage. This scoring was applied independently by two of the authors, with disagreements resolved by discussion; we note that this process is a lightweight quality screen rather than a formal risk-of-bias tool such as those used in clinical systematic reviews (e.g., Cochrane RoB), and we identify the adoption of a more formal risk-of-bias framework as an item for future work (Section VI.D).

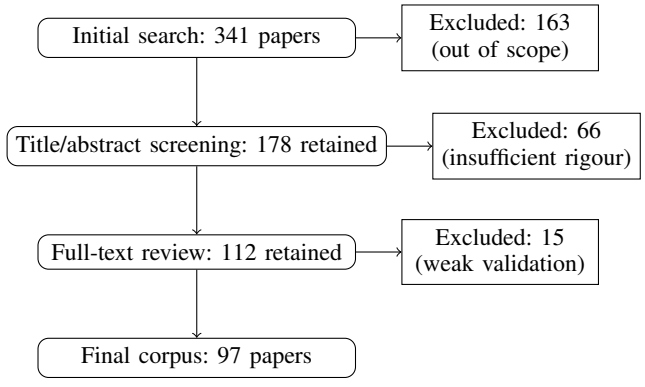

Fig. 1. PRISMA-style literature selection workflow, from 341 candidate papers to a final corpus of 97 included studies, with a lightweight quality assessment applied at the final stage.

III. Literature Review

A. AI and Machine Learning for Threat Detection

A recurring theme in the reviewed literature is that traditional rule-based SIEM systems struggle to keep pace with the volume and sophistication of modern threats [14], [26]. These systems generate large numbers of alerts, many of which are false positives, contributing to analyst fatigue and, in some cases, degraded detection quality [13], [31].

Several deep learning approaches have been explored as alternatives or complements to rule-based detection. Convolutional Neural Networks (CNNs) have been applied to identify patterns and signatures of malicious behaviour in high-dimensional traffic and file data [9], [15]. Recurrent architectures, particularly Long Short-Term Memory (LSTM) networks, have been used for security data with sequential or time-series characteristics, including detection of Advanced Persistent Threats (APTs) that unfold over extended periods [26]. Generative Adversarial Networks (GANs) have been used both to synthesise anomaly examples for training and to study adversarial robustness of detection models [17]. Where data-sharing constraints exist, Federated Learning has been proposed as a means for organisations to collaboratively train threat-intelligence models without exposing raw data [10], [30].

B. Blockchain for Forensic Integrity

Blockchain technology has attracted interest as a mechanism for preserving the integrity of digital forensic evidence [22], [23]. Distributed ledger platforms, particularly Hyperledger Fabric, have been proposed to provide an immutable audit trail: once an event record is committed, it cannot be silently altered by an external attacker or a malicious insider [34], [35]. Several studies further propose the use of smart contracts to automatically capture and seal evidence at the moment it is generated, reducing reliance on manual chain-of-custody procedures [36].

C. Automated Response and Orchestration

Security Orchestration, Automation, and Response (SOAR) platforms have been studied as a means of reducing the manual workload associated with repetitive incident-response tasks [8], [12]. A number of studies explore the use of Reinforcement Learning (RL) to allow response policies to adapt over time based on observed outcomes [11], [31], although these approaches are generally reported to require substantially more training time than static, predefined playbooks. In several deployments, response actions are implemented through configuration-management tools such as Ansible, which execute remediation playbooks across heterogeneous infrastructure [49].

D. Multi-Cloud Security Monitoring

As organisations increasingly deploy services across multiple cloud providers, monitoring has become more fragmented. Major providers—AWS CloudTrail, Azure Monitor, and Google Cloud Security Command Center—each offer native telemetry, but correlating events across these platforms remains an open challenge in practice [19], [20]. Containerised environments orchestrated with Kubernetes introduce a further layer of complexity for monitoring and correlation [13].

E. Human Factors in SOC Operations

The human dimension of SOC work is consistently identified as an underaddressed factor in SOC effectiveness. Studies report that analyst decision quality tends to degrade during sustained periods of high alert volume [45], and that wargaming-based training exercises can improve team preparedness for real incidents [41]. The concept of a "Cognitive SOC"—an approach intended to reduce analyst cognitive load through smarter alert triage and contextual enrichment—has been proposed as an early-stage but promising direction [44].

IV. Proposed Conceptual Framework: CDSOCMP

A. Framework Overview and Status

Building on the gaps identified in Section III, we propose the Cyber Defence Security Operations Centre Management Platform (CDSOCMP) as a **conceptual** four-phase architecture intended to address the four recurring SOC problems in a coordinated design. We reiterate that CDSOCMP is a design proposal synthesised from the reviewed literature; it has not been implemented, deployed, or empirically tested, and the architecture should be read as a set of design goals and component choices rather than a validated system. Conceptually, CDSOCMP is intended to provide:

1) AI-based detection and automated response designed to be tightly integrated from the outset, rather than combined as separate add-on tools.
2) Forensic evidence anchored to a blockchain at the time of detection, rather than logged retrospectively.
3) A single monitoring view spanning AWS, Azure, and GCP.
4) A forensic design intended to reduce opportunities for evidence tampering, rather than relying solely on post-incident recovery.

B. The Four-Phase Architecture

The proposed architecture, shown in Fig. 2, consists of four sequential phases connected by a feedback loop.

Phase 1 – Data Ingestion. Telemetry is intended to be drawn from heterogeneous sources, including AWS CloudTrail, Azure Monitor, GCP Security Command Center, syslog/CEF streams, and endpoint agents. A normalisation layer would convert these vendor-specific formats into a common event schema for use by later phases.

Phase 2 – AI Processing. Normalised events would be analysed by a hybrid CNN-GAN detection component: the CNN component for known-signature and behavioural-anomaly detection, and the GAN component for generating synthetic adversarial examples intended to stress-test detection thresholds during development and retraining. An Explainable AI (XAI) layer, for example using SHAP values [12], is

intended to accompany each detection with a human-readable rationale to support analyst review.

Phase 3 – Blockchain Anchoring. Detections exceeding a configurable confidence threshold would be cryptographically hashed and recorded on a Hyperledger Fabric ledger at the time of detection, with the aim of producing a tamper-evident, time-stamped chain of custody from the moment of discovery rather than after the fact.

Phase 4 – Automation Trigger. Confidence-scored detections would be mapped to pre-defined Ansible playbooks corresponding to threat type and severity, with possible actions including network node isolation, firewall rule changes, credential revocation, or escalation to a human analyst for ambiguous cases. All automated actions in this phase would also be recorded on the Phase 3 ledger.

A continuous feedback loop is intended to connect outcomes from Phase 4 back to Phase 2, allowing detection models to be retrained using outcome data over time. Table II summarises the technologies envisaged for each phase.

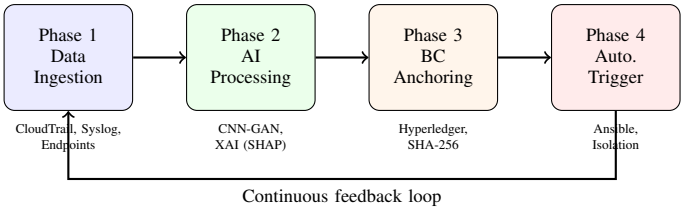

Fig. 2. Proposed CDSOCMP four-phase conceptual architecture, from Data Ingestion through AI Processing and Blockchain Anchoring to the Automation Trigger, with a continuous feedback loop. This diagram represents a design proposal and not an implemented system.

C. Anticipated Trade-offs and Computational Considerations

Combining a hybrid CNN-GAN detector, an LSTM classifier, SHAP-based explanation generation, blockchain anchoring, and SOAR-style automation into a single pipeline introduces non-trivial computational and operational overhead, which existing literature on individual components suggests but does not, in combination, fully quantify. In particular: (i) GAN-based components are reported elsewhere to be unstable to train and may require careful retraining schedules [17]; (ii) per-event blockchain anchoring introduces transaction throughput and latency considerations that scale with telemetry volume, and Hyperledger Fabric's permissioned consensus would need to be sized accordingly; (iii) SHAP-based explanation generation adds inference-time overhead proportional to model complexity and input dimensionality; and (iv) RL-based response adaptation, where used, is reported in the literature to converge slowly [11], [31]. We therefore identify the computational cost of operating this combined pipeline at production telemetry volumes as a key open question for prototype evaluation (Section VI.D), rather than a property that can be estimated from the individually reported figures of its component technologies.

TABLE II
CDSOCMP CONCEPTUAL ARCHITECTURE: PHASE DETAILS

Phase	Component	Intended Function
1: Ingestion	AWS CloudTrail	Cloud API audit logging
	Syslog / CEF	Host and network telemetry
	Endpoint agents	Process and file monitoring
2: AI Proc.	Hybrid CNN-GAN	Signature + anomaly detection
	LSTM classifier	Sequential threat detection
	XAI (SHAP)	Analyst-readable explanations
3: BC Anchor	Hyperledger Fabric	Distributed ledger
	SHA-256 hashing	Event integrity verification
	Smart contracts	Automated evidence sealing
4: Auto. Resp.	Ansible playbooks	Orchestrated remediation
	Node isolation	Lateral movement containment
	Escalation rules	Human-in-the-loop oversight

V. ANALYSIS OF THE REVIEWED LITERATURE

A. Paper Categorisation

Table III shows the distribution of the 97 reviewed papers across five research categories, based on each paper's primary contribution as identified during full-text review. AI/ML threat detection represents the largest category (33%), reflecting substantial academic interest in applying deep learning to security detection tasks. SOC operations and human factors form the second-largest category (25%), which we read as an indication that a portion of the research community views analyst workload and decision-making as significant open problems independent of detection accuracy.

TABLE III
DISTRIBUTION OF REVIEWED PAPERS BY RESEARCH CATEGORY

Category	Count	Proportion
AI/ML Threat Detection	32	33%
SOC Ops & Human Factors	24	25%
Automated Response (SOAR)	15	15%
Blockchain Security	12	12%
Multi-Cloud Monitoring	8	8%
Integrated AI–Blockchain	6	6%
Total	**97**	**100%**

B. Comparison of AI Techniques for Threat Detection

Table IV summarises the main AI/ML techniques identified in the reviewed literature, together with the strengths and limitations reported for SOC-related detection tasks.

TABLE IV
COMPARATIVE ANALYSIS OF AI TECHNIQUES FOR THREAT DETECTION (AS REPORTED IN THE REVIEWED LITERATURE)

Technique	Reported Strengths	Reported Limitations
CNN	High-accuracy pattern recognition in high-dimensional data	Requires large labelled datasets; weak at temporal reasoning
RNN/LSTM	Effective on sequential and time-series security events	Computationally heavy; vanishing gradients at scale
GAN	Generates synthetic anomalies to augment training data	Training instability; mode collapse in adversarial settings
Federated Learning	Privacy-preserving cross-organisational intelligence sharing	Communication overhead; uneven data quality across participants
Reinforcement Learning	Adaptive, self-improving response strategies	Slow convergence; sparse reward signals
Graph Neural Network	Analyses entity relationships and lateral movement	Scales poorly on large enterprise graphs

C. Reported Detection Accuracy Across Studies

Fig. 3 summarises detection accuracy figures as reported by individual studies in the reviewed corpus for each AI technique [9]–[11], [15], [17], [26]. **We emphasise an important caveat regarding this figure: the values shown were obtained by different research teams, using different datasets (e.g., NSL-KDD, CICIDS, proprietary enterprise traffic), different evaluation protocols, and different threat scenarios.** They are not directly comparable in a statistical sense, and the figure should be read as an indicative summary of the *range* of accuracy values reported in the literature for each technique class, not as a controlled benchmark. In particular, the hybrid CNN-GAN figure of approximately 97% reflects the highest value reported among the studies reviewed for that technique combination [9], [15], [17]; it is *not* a result obtained for CDSOCMP, and no claim is made that CDSOCMP would achieve this figure if implemented.

D. MTTD/MTTR Improvements Reported for Existing Architecture Classes

Fig. 4 summarises Mean Time to Detect (MTTD) and Mean Time to Respond (MTTR) improvement percentages reported in the literature for three existing architecture classes: rule-based SIEM, AI-only detection, and SOAR-only automation [1], [8], [11], [49]. These figures are drawn from the reviewed primary studies, each of which used its own baseline, dataset, and deployment context, and are summarised here purely to characterise the *range* of improvements that have been reported for single-technology approaches in the literature.

We do not report an MTTD/MTTR figure for CDSOCMP in this figure or elsewhere in this paper, as CDSOCMP has not been implemented or evaluated. Based on the architectural rationale in Section IV—namely, that CDSOCMP is designed to integrate detection, blockchain anchoring, and automated response within a single pipeline rather

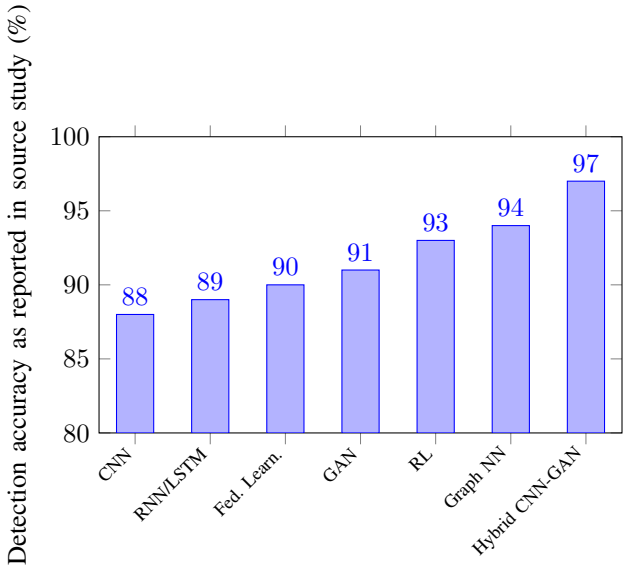

Fig. 3. Highest detection accuracy reported by individual studies for each AI technique class, as found in the reviewed literature. Values originate from heterogeneous datasets and evaluation protocols and are not directly comparable; they are not results for CDSOCMP. See Table IV for the corresponding source studies.

than as separate tools—we hypothesise that an integrated approach *could* yield improvements at or above the upper end of the ranges reported for single-technology approaches. This is presented explicitly as a hypothesis to be tested through prototype evaluation (Section VI.D), not as a measured or projected result.

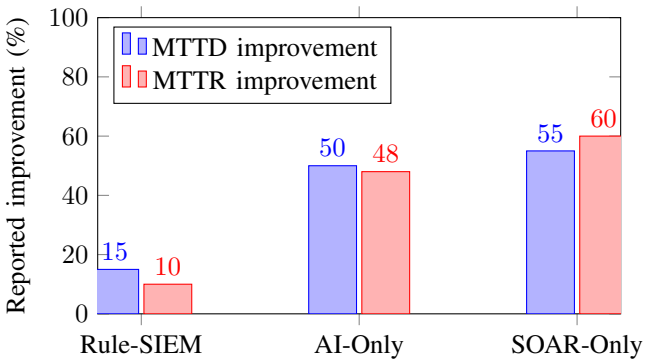

Fig. 4. MTTD and MTTR improvement percentages reported in the literature for existing single-technology SOC architecture classes, relative to each study's own baseline [1], [8], [11], [49]. No CDSOCMP figure is shown, as CDSOCMP has not been implemented or evaluated.

E. SOC Technology Evolution

Fig. 5 summarises the three broad generations of SOC technology discussed in Section I, ending with the research-stage concepts – integrating AI, orchestration, and blockchain – that CDSOCMP is intended to extend.

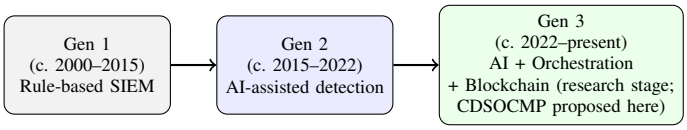

Fig. 5. Broad generations of SOC technology evolution as characterised in the reviewed literature, culminating in the research-stage integrated concepts that motivate the CDSOCMP proposal.

VI. Discussion

A. Key Findings from the Review

Four observations emerge from the literature synthesis:

First, hybrid AI architectures – particularly CNN-GAN combinations – are reported in individual studies to outperform single-model architectures, with the highest reported figures around 97% versus 88–94% for individual techniques in the corresponding studies [9], [15], [17]. As noted in Section V.C, these figures come from heterogeneous evaluation setups and should not be read as a controlled comparison.

Second, real-time blockchain anchoring of forensic evidence – as opposed to retrospective logging – appears to be comparatively underexplored in the literature relative to AI-based detection, representing a gap that CDSOCMP's Phase 3 is intended to address [22], [36].

Third, Reinforcement Learning-based SOAR approaches are reported as a promising direction for adaptive response, but multiple studies note substantially longer training times compared to static playbook-based approaches [11], [31].

Fourth, the human dimension of SOC work – particularly analyst cognitive load – is comparatively underrepresented in the technical literature relative to its reported impact on detection outcomes [44], [45]. We note, however, that the quantitative evidence base on this point is itself limited: the studies we reviewed on this topic [44], [45] are largely qualitative or survey-based, and we did not identify large-scale quantitative studies directly linking analyst cognitive load to detection or response outcomes. We treat this as an open empirical question rather than an established quantitative relationship, and flag it as a candidate area for future controlled study.

B. Comparison with Commercial SOC Platforms

Several commercial and open-source platforms already address parts of the problem space discussed in this review. Microsoft Sentinel and IBM QRadar provide AI-assisted SIEM capabilities with built-in anomaly detection and integration across cloud and on-premises sources. Splunk SOAR and Palo Alto Cortex XSOAR provide mature playbook-based automation and case-management capabilities, with marketplace integrations across a wide range of security tools. None of these platforms, to our knowledge, natively anchor individual detection events to a permissioned blockchain ledger at the time of detection; where blockchain-based logging has been explored commercially, it has generally been positioned as an add-on for compliance archiving rather than as a real-time component of the detection-to-response pipeline. CDSOCMP's Phase 3 (real-time blockchain anchoring, tightly coupled to Phases 2 and 4) is therefore intended as a complement to – rather than a wholesale replacement for – the detection and orchestration capabilities already mature in platforms such as Sentinel, QRadar, Splunk SOAR, and Cortex XSOAR. A direct empirical comparison between CDSOCMP and these commercial platforms would require a working prototype and is identified as a priority for future work (Section VI.D).

C. Adversarial Attacks Against AI-Driven SOCs

The reliance of CDSOCMP's Phase 2 on deep learning models for detection introduces exposure to adversarial machine learning attacks, a topic addressed by a subset of the reviewed literature [17]. Reported attack categories relevant to SOC detection models include evasion attacks, in which inputs are perturbed to avoid triggering a detection model while preserving malicious functionality; poisoning attacks, in which an adversary influences training data (a particular concern for the Federated Learning approaches discussed in Section III.A); and model-extraction attacks, in which an adversary probes a deployed model to reconstruct its decision boundaries. The GAN component proposed within CDSOCMP's Phase 2 is intended, in part, to improve robustness against evasion-style attacks by exposing the CNN/LSTM detectors to synthetic adversarial examples during training; however, GAN-based adversarial training is itself reported to be unstable and prone to mode collapse [17], and we do not claim that this design choice resolves the adversarial robustness problem. We identify formal adversarial robustness evaluation of any CDSOCMP prototype as a necessary component of future validation work.

D. Limitations and Future Work

This review and the accompanying CDSOCMP proposal have several limitations that should be considered when interpreting the contributions of this paper.

First, and most importantly, **no live or production SOC has been used to test CDSOCMP**, and no prototype currently exists. All performance figures discussed in this paper (Figs. 3 and 4) are drawn from the reviewed component-level studies and characterise those studies' own reported results, not the performance of CDSOCMP as an integrated system. Any expectation of how an integrated CDSOCMP system would perform remains a hypothesis pending prototype evaluation.

Second, the quality assessment described in Section II.C is a lightweight, two-criteria-reviewer screening process rather than a formal risk-of-bias framework. Adopting a more established framework – for example, an adaptation of the Cochrane risk-of-bias approach or a software-engineering-oriented checklist such as those used in systematic literature reviews of software engineering research – is identified as a priority for any extended version of this review.

Third, the computational overhead of operating a combined CNN-GAN-LSTM-SHAP-blockchain-SOAR pipeline at production telemetry volumes has not been characterised (Section IV.C) and represents a significant open question, partic-

ularly given that several of the constituent techniques (GAN training, per-event blockchain anchoring, SHAP computation) are individually reported to be resource-intensive.

We identify three priorities for future work. First, develop a prototype implementation of CDSOCMP and evaluate it under controlled conditions using a recognised intrusion-detection dataset and a representative multi-cloud testbed, including direct comparison against representative commercial platforms (Section VI.B). Second, extend the architecture to incorporate Zero Trust Architecture (ZTA) principles [51]; in particular, Phase 4's automated response actions (node isolation, credential revocation) align naturally with ZTA's continuous-verification model, and we see an opportunity to use ZTA policy decision points as an additional input to Phase 4's confidence-scoring logic. Third, formally incorporate Cognitive SOC principles [44] into the analyst-facing components of Phase 2's XAI output, with the goal of reducing analyst cognitive load – a goal that should itself be evaluated quantitatively as part of prototype testing, given the limitation noted in Section VI.A regarding the current evidence base on this topic.

VII. Conclusion

This paper has presented a systematic review of 97 studies, published between 2013 and 2026, examining the application of AI and blockchain technologies to four persistent SOC challenges: fragmented visibility, alert fatigue, slow incident response, and forensic vulnerability. The review confirms findings consistent with prior literature: AI-based techniques are widely reported to improve detection accuracy relative to rule-based approaches; blockchain-based approaches are proposed as a means of preserving forensic integrity; and automation is widely reported to reduce response times relative to manual processes. However, the review also confirms that no single platform in the existing literature integrates all four elements within one coordinated architecture.

In response to this gap, we have proposed CDSOCMP, a four-phase *conceptual* architecture comprising AI-driven detection, real-time blockchain anchoring, confidence-triggered automated response, and unified multi-cloud monitoring. We have deliberately refrained from reporting performance figures for CDSOCMP itself, as no implementation or evaluation currently exists; the accuracy and improvement figures discussed in this paper characterise the existing literature on individual component technologies and should not be interpreted as CDSOCMP results. The principal next step, and the central priority for future work, is the development of a working prototype and its evaluation under controlled, reproducible conditions, including direct comparison with commercial SOC platforms and assessment of adversarial robustness.

Acknowledgment

The authors would like to thank the Department of Information and Communication Engineering at The Islamia University of Bahawalpur for their support.

References

[1] A. Sharma and R. Spunda, "SOC optimisation through AI-powered automation and blockchain integration," *J. Cybersecurity Technol.*, 2025.
[2] O. S. Ndibe, "AI-driven forensic systems for real-time anomaly detection and threat mitigation," *IEEE Access*, 2025.
[3] P. Shelke and T. Hämäläinen, "Analysing multidimensional strategies for cyber threat detection," *Computers & Security*, 2024.
[4] L. Tan et al., "AI-based intrusion detection system for software-defined networks," *IEEE Trans. Netw. Serv. Manag.*, 2024.
[5] J. Zhang et al., "Deep learning for cyber security: A comprehensive survey," *Neurocomputing*, 2024.
[6] A. Alzahrani and D. Alghazzawi, "Intrusion detection in Internet of Things using deep learning," *IEEE Internet Things J.*, 2023.
[7] M. Abomhara et al., "Machine learning for cybersecurity: A systematic review," *ACM Comput. Surv.*, 2023.
[8] J. Rodriguez et al., "AI-based security in 5G networks: A survey," *IEEE Commun. Surv. Tutor.*, 2023.
[9] J. Li et al., "Deep learning-based malware detection for Android," *Comput. Secur.*, 2023.
[10] Y. Zhao et al., "Federated learning for privacy-preserving cyber threat intelligence," *IEEE Trans. Inf. Forensics Secur.*, 2023.
[11] N. Standen et al., "Reinforcement learning for autonomous cyber defence," *ACM Comput. Surv.*, 2023.
[12] M. Gupta et al., "Explainable AI for network security: A SHAP-based approach," *Comput. Netw.*, 2023.
[13] K. Patel et al., "AI-driven vulnerability assessment in cloud-native applications," *IEEE Cloud Comput.*, 2023.
[14] M. Al-Sarem et al., "Detecting phishing websites using hybrid deep learning models," *IEEE Access*, 2023.
[15] S. Wang et al., "Graph neural networks for botnet detection in IoT," *IEEE Internet Things J.*, 2023.
[16] X. Liu et al., "AI-driven security for smart power grids," *IEEE Trans. Smart Grid*, 2023.
[17] X. Yuan et al., "Adversarial attacks and defences in AI-based cybersecurity," *IEEE Trans. Inf. Forensics Secur.*, 2023.
[18] A. Singh et al., "AI-driven threat hunting in large-scale enterprises," *Comput. Secur.*, 2023.
[19] K. Sundararajan et al., "Deep learning for biometric security: A review," *Pattern Recognit.*, 2023.
[20] Z. El-Rewini et al., "AI-based security for autonomous vehicles," *IEEE Trans. Intell. Transp. Syst.*, 2023.
[21] P. Shelke and T. Hämäläinen, "Analysing the role of AI in improving network security monitoring," *IEEE Netw.*, 2024.
[22] R. Garcia et al., "Robust malware classification using convolutional neural networks," *IEEE Trans. Inf. Forensics Secur.*, 2024.
[23] H. Wang et al., "Securing smart grids with AI-based anomaly detection," *IEEE Trans. Smart Grid*, 2024.
[24] A. Kumar, "AI-powered anti-phishing tools for email security," *Comput. Secur.*, 2024.
[25] D. Wilson, "Using generative AI (LLMs) to write secure code," *IEEE Secur. Privacy*, 2024.
[26] J. Li and Y. Zhang, "Advanced persistent threat (APT) detection using LSTM," *Comput. Secur.*, 2023.
[27] A. Alzahrani et al., "A review of deep learning techniques for Android malware detection," *IEEE Access*, 2023.
[28] Z. Ahmad et al., "Deep learning for network intrusion detection: A systematic review," *Comput. Secur.*, 2023.
[29] J. Rodriguez et al., "AI-based security management in 5G networks," *IEEE Commun. Mag.*, 2023.
[30] Y. Zhao et al., "Federated learning for privacy-preserving cyber threat intelligence," *IEEE Trans. Inf. Forensics Secur.*, 2023.
[31] N. Standen et al., "Reinforcement learning for autonomous cyber defence," *ACM Comput. Surv.*, 2023.
[32] M. Gupta et al., "Explainable AI for network security: A SHAP-based approach," *Comput. Netw.*, 2023.
[33] K. Patel et al., "AI-driven vulnerability assessment in cloud-native applications," *IEEE Cloud Comput.*, 2023.
[34] M. Al-Sarem et al., "Detecting phishing websites using hybrid deep learning models," *IEEE Access*, 2023.
[35] Y. Xin et al., "Deep learning for cyber security: A review and case studies," *IEEE Access*, 2023.
[36] L. Caviglione et al., "AI-powered digital forensics: Challenges and future directions," *Digit. Investig.*, 2023.

[37] S. Wang et al., “Graph neural networks for botnet detection in IoT,” *IEEE Internet Things J.*, 2023.
[38] X. Liu et al., “AI-driven security for smart power grids,” *IEEE Trans. Smart Grid*, 2023.
[39] K. Gibert et al., “Deep learning for malware detection: A survey,” *Comput. Secur.*, 2023.
[40] T. Tang et al., “AI-based intrusion detection in software-defined networks,” *IEEE Trans. Netw. Serv. Manag.*, 2023.
[41] P. Sarjakivi et al., “Using wargaming to model cyber defence decision-making,” in *Proc. IEEE Conf. Cyber Secur.*, 2024.
[42] F. Jalalvand et al., “Alert prioritisation in SOCs: A systematic survey,” *ACM Comput. Surv.*, 2024.
[43] M. Wurzenberger et al., “NEWSROOM: Towards automating cyber situational awareness,” *Comput. Secur.*, 2024.
[44] F. Binbeshr et al., “The rise of cognitive SOCs: A systematic literature review,” *IEEE Access*, 2025.
[45] A. Reeves and D. Ashenden, “Understanding decision making in SOCs,” *Hum. Factors*, 2023.
[46] D. Ball and G. Waters, *Cyber Defence and Warfare*. Canberra: ANU Press, 2013.
[47] M. Abd Majid et al., “Model for the successful development and implementation of SOC,” *IEEE Access*, 2021.
[48] C. Onwubiko and K. Ouazzane, “Challenges towards building an effective SOC,” *J. Cyber Secur. Technol.*, 2022.
[49] S. Vyas et al., “Automated cyber defence: A review,” *IEEE Access*, 2023.
[50] K. J. Ferguson-Walter, “An empirical assessment of the effectiveness of deception for cyber defence,” *IEEE Trans. Inf. Forensics Secur.*, 2020.
[51] M. L. Gambo and A. Almulhem, “Zero trust architecture: A systematic literature review,” *IEEE Access*, 2026.
[52] Q. Shah and S. Aravazhi, “AI and blockchain integration for next-generation cyber security,” *Comput. Secur.*, 2024.